Long-term implementation of the cooperative solution in a multistage multicriteria game

Denis Kuzyutin[a,b], Nadezhda Smirnova[b,a,*], Ekaterina Gromova[a]

[a] *Saint Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, 199034, Russia*
[b] *National Research University Higher School of Economics (HSE), Soyuza Pechatnikov ul. 16, St. Petersburg, 190008, Russia*

Abstract

In order to find an optimal and time consistent cooperative path in multicriteria multistage game the minimal sum of relative deviations rule is introduced. Using this rule one can construct a vector-valued characteristic function that is weakly superadditive. The sustainability of the cooperative agreement is ensured by using an imputation distribution procedure (IDP) based approach.

We formulate the conditions an IDP should satisfy to guarantee that the core is strongly time consistent (STC). Namely, if the imputation distribution procedure for the Shapley value satisfies the efficiency condition, the strict balance condition and the strong irrational-behavior-proof condition, given that the Shapley value belongs to the core of each subgame along the cooperative path, it can be used as a "supporting imputation" which guarantees that the whole core is STC. We discuss three payment schedules and check whether they can be used as supporting imputation distribution procedures for the considered multicriteria game.

Keywords: Dynamic game, multiple criteria decision making, multicriteria game, strong time consistency, Shapley value, cooperative solution

1. Introduction

The theory of multicriteria games (multiobjective games or games with vector payoffs) lies at the intersection of Multiple Criteria Decision Making (MCDM) and classical game theory. It can be used to model various real-world situations where several decision makers (or players) need to consider several goals when choosing their strategies (see, e.g., [41, 2, 44, 22, 5, 20, 19, 21]). As noted in [3], one of the important trends in MCDM is "the development of adequate dynamic MCDM approaches, taking into account the influence of time in evolving decision processes". This paper deals with the dynamic properties of cooperative behavior in n-person multicriteria multistage games with perfect information [24, 30, 16] and hence falls in with the specified research direction.

Different cooperative solutions for static and dynamic games with vector payoffs were studied in [1, 11, 5, 4, 33, 34, 17, 18, 38, 36, 15]. In order to achieve and implement a long-term cooperative agreement in multicriteria dynamic games we have to solve the following problems. First, when players seek to achieve the maximal total vector payoff of the grand coalition, they face the problem of choosing a unique Pareto efficient payoffs vector. In the dynamic setting it is necessary that a specific method the players agreed to accept in order to select a particular Pareto efficient solution not only takes into account the relative importance of the criteria, but also satisfies time consistency (see, e.g., [27, 10, 30]). That is to say, a fragment of the optimal cooperative trajectory in the subgame should remain optimal in this subgame. For the special case when the criteria have significantly different importance, and all players arrange the criteria in the same order the refined leximin algorithm introduced in [15] is a reasonable approach to find a time consistent cooperative path. Otherwise, say, when the players rank the criteria in a different order or some of the criteria have approximately equal importance, the players need to employ other appropriate methods to select a unique Pareto efficient solution (see, e.g., [35, 9]). In this paper, we use the rule of minimal sum of relative deviations (MSRD) from the ideal payoffs vector [23] to find a unique optimal cooperative path, and prove its time consistency.

After choosing the cooperative trajectory it is necessary to construct a vector-valued characteristic function. To this end, we suggest to use the ζ-characteristic function introduced in [7] and the MSRD rule in order to select a particular Pareto efficient solution for the auxiliary vector optimization problems. We assume that the cooperative multicriteria game satisfies the component-wise transferable utility property, i.e., we allow the payoff to be transferred between the players within the same criterion k. Note that the main measurable criteria used in multicriteria resource management problems (for instance, water

*E-mail address of corresponding author nadezhda.v.smirnova@gmail.com

supply amount, cost of water purification, the quota for the use of a common resource, the quota for emissions, profit from the use of a common resource, etc.) satisfy the required transferable utility property. To determine the optimal payoff allocation we explore the vector analogues of two well known cooperative solutions, namely the core and the Shapley value [1, 34, 33].

Lastly, to guarantee the sustainability of the achieved long-term cooperative agreement one needs to design a consistent imputation distribution procedure or a payment schedule [28, 31, 30, 37] that should satisfy a set of useful properties. The detailed review of dynamic properties the IDP may satisfy for multistage multicriteria games is presented in [17, 18, 15]. In this paper we mainly focus on the efficiency constraint and the strict balance condition as well as strong time consistency (STC) [30, 29, 6, 32, 39] and strong irrational-behavior-proofness (IBP) [45] for individual players and coalitions. Although the criteria are measured in different units, we use the word *payment* with respect to every criterion k for the sake of simplicity and uniformity. It is worth noting that the IDP-based approach proved to be an effective method to implement long-term cooperative agreement in single-criterion dynamic games (see, e.g., [28, 30, 31, 37]).

In the paper, we establish the exact set of properties that an imputation distribution procedure has to satisfy in order to guarantee the strong time consistency of the core. Namely, if the Shapley value (or any other single-valued cooperative solution) belongs to the core at every subgame along the optimal cooperative path, and the payment schedule satisfies the efficiency, strict balance and strong irrational-behavior-proof conditions, the core in multicriteria multistage game satisfies STC if the Shapley value is used as the supporting imputation. We consider three payment schedules which satisfy the efficiency and strict balance condition and check the STC of the core for given 3-person 3-criteria multistage game.

There are a number of possible applications of the proposed method. In particular, it can be used when analyzing water resources management problems which are often modeled as dynamic MCDM problems (see [20] for the review of water resources conflict resolution models). For instance, in [21], a multicriteria game theoretic approach is applied to analyze California's Sacramento-San Joaquin Delta problem [19], and "the most likely" Pareto efficient cooperative solution (building a tunnel or pipeline to move water around the Delta) is proposed. In this paper we examine an alternative approach to select optimal cooperative solution, and provide a step-by-step method to implement the long-term cooperative agreement which guarantees the sustainability of cooperation.

The contribution of this paper is twofold:

- we suggest the minimal sum of relative deviations rule as a specific method to find an optimal and time consistent cooperative trajectory for multicriteria multistage game and to construct a vector-valued characteristic function. This method for sectioning a unique Pareto efficient solution differs from the one used in [15] and is applicable to a broader class of MCDM problems. Note that the constructed ζ-characteristic function is proved to satisfy weak superadditivity property.

- we provide a general characterization of the strong time consistent core in multicriteria multistage games which allows the players to use any appropriate imputation distribution procedure along the cooperative trajectory. This is a generalization of the Propositon 3, proved in [15] for the partial case of the so-called incremental payment schedule. We propose an example to clarify how the players can select an appropriate imputation distribution procedure on the base of their properties analysis.

The paper is organized as follows: The class of finite multistage r-criteria games in extensive form with perfect information is formalized in Section 2. In Section 3.1, we prove that the set of all Pareto efficient strategy profiles does not satisfy the STC property. In Section 3.2, the minimal sum of relative deviations rule for choosing optimal cooperative trajectory is formalized and the resulting Pareto optimal solution is proved to be time consistent. A vector-valued ζ-characteristic function for a multicriteria cooperative game based on the minimal sum of relative deviations rule is constructed and proved to be weakly superadditive in Section 4. In Section 5.1, we specify which properties the payment schedule should satisfy to guarantee sustainable cooperation in dynamic multicriteria game. The general conditions for STC of the core are proved in Section 5.2. We examine an illustrative example and compare three different payment schedules in Section 6. Finally, Section 7 presents a brief conclusion.

2. Multistage multicriteria game with perfect information

We consider a finite multistage multicriteria game (game in extensive form) with perfect information comprised by the following ingredients (see [12, 30, 16, 17] for details):

- $N = \{1, ..., n\}$ is the set of all players.
- K is the (rooted) game tree with the root x_0 and the set of all nodes P.
- $S(x)$ is the set of all direct successors of the node x, and $S^{-1}(x)$ denotes the unique predecessor of the node $x \neq x_0$ such that $x \in S(S^{-1}(x))$.
- P_i is the set of all decision nodes of the i-th player (at these nodes the player i chooses the following node), $P_i \cap P_j = \varnothing$, for all $i, j \in N$, $i \neq j$, and $P_{n+1} = \{z^j\}_{j=1}^m$ denotes the set of all terminal nodes (final positions), $S(z^j) = \varnothing\ \forall z^j \in P_{n+1}$. It holds that $\cup_{i=1}^{n+1} P_i = P$.
- $\omega = (x_0, \ldots, x_{t-1}, x_t, \ldots, x_T)$ is the path (or the trajectory) in the game tree, $x_{t-1} = S^{-1}(x_t), 1 \leqslant t \leqslant T$;

$x_T = z^j \in P_{n+1}$; where index t in x_t denotes the ordinal number of this node in the path ω (in discrete time).

- $h_i(x) = (h_{i/1}(x), \dots, h_{i/r}(x))$ is the r-component vector payoff of the player i computed at the node $x \in P \backslash \{x_0\}$. We assume that for all $i \in N$, $k = 1, ..., r$, and $x \in P \backslash \{x_0\}$ the respective payoffs are positive, i.e., $h_{i/k}(x) > 0$.

In the following, we will write Γ^{x_0} when referring to the multistage multicriteria game defined above. Since we deal with the multistage games with perfect information we consider only the class of players' pure strategies (see [12, 24, 30] for details). The pure strategy $u_i(\cdot)$ of the i-th player is a function that uniquely determines for each node $x \in P_i$ the next node $u_i(x) \in S(x)$ that the player i has to choose at x. Denote by U_i the set of all possible pure strategies of the player i , and $U = \prod_{i \in N} U_i$. Every pure strategy profile $u = (u_1, ..., u_n) \in U$ generates the unique path $\omega(u) = (x_0, x_1(u), \dots, x_t(u), x_{t+1}(u), \dots, x_T(u))$, where $x_{t+1} = u_j(x_t) \in S(x_t)$ if $x_t \in P_j$, $0 \le t \le T-1$.

Furthermore, each path $\omega(u)$ generates a collection of the vector payoffs of all players. We will write

$$H_i(u) = (H_{i|1}(u), \dots, H_{i|r}(u)) = \sum_{\tau=1}^{T} h_i(x_\tau(u)), \qquad (1)$$

to denote the value of the i-th player's vector payoff function which corresponds to the strategy profile $u = (u_1, ..., u_n)$.

Following [12, 30], at every intermediate node $x_t \in P \backslash P_{n+1}$ in the game Γ^{x_0} one can define a subgame Γ^{x_t} with the subgame tree K^{x_t} and the subroot x_t and a factor-game Γ^D with the factor-game tree $K^D = \{x_t\} \cup (K \backslash K^{x_t})$. Decomposition of the original multistage game Γ^{x_0} at the intermediate node x_t onto the subgame Γ^{x_t} and the factor-game Γ^D further induces the corresponding decomposition of (pure and mixed) strategies (see [12, 30] for detailes).

Denote by $P_i^{x_t}(P_i^D)$, $i = 1, \dots, n$, the restriction of the set P_i onto the subgame tree $K^{x_t}(K^D)$, and let $u_i^{x_t}(u_i^D)$ denote the corresponding restriction of the i-th player's pure strategy $u_i(\cdot)$ in Γ^{x_0} to $P_i^{x_t}(P_i^D)$. The strategy profile $u^{x_t} = (u_1^{x_t}, \dots, u_n^{x_t})$ in the subgame generates the path (trajectory) $\omega^{x_t}(u^{x_t}) = (x_t, x_{t+1}(u^{x_t}), \dots, x_T(u^{x_t}))$ and, therefore, a collection of all the players' vector payoffs in this subgame. Similarly to (1), we will denote by

$$H_i^{x_t}(u^{x_t}) = \tilde{h}_i^{x_t}(\omega^{x_t}(u^{x_t})) = \sum_{\tau=t+1}^{T} h_i(x_\tau(u^{x_t})), \qquad (2)$$

the value of the i-th player's vector payoff function in Γ^{x_t}, and by $U_i^{x_t}$ the set of all possible i-th player's pure strategies in the subgame Γ^{x_t}, $U^{x_t} = \prod_{i \in N} U_i^{x_t}$. Moreover,

$$\begin{aligned} H_i(u) = &\sum_{\tau=1}^{t} h_i(x_\tau(u)) + \sum_{\tau=t+1}^{T} h_i(x_\tau(u^{x_t})) = \\ &= \tilde{h}_i(\underline{\omega}^{x_t}(u)) + \tilde{h}_i^{x_t}(\omega^{x_t}(u^{x_t})), \end{aligned} \qquad (3)$$

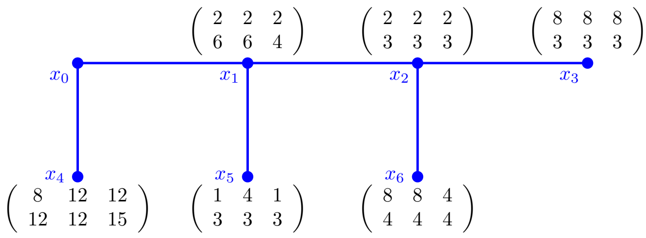

Figure 1: The bicriteria game tree

where $\underline{\omega}^{x_t}(u) = (x_0, x_1, \dots, x_{t-1}, x_t)$ denotes a fragment of path $\omega(u)$ implemented before the start of the subgame Γ^{x_t}.

Note that, since $P_i = P_i^{x_t} \cup P_i^D$, we can compose the i-th player's pure strategy $W_i = (u_i^D, v_i^{x_t}) \in U_i$ in the original game Γ^{x_0} from her strategies $v_i^{x_t} \in U_i^{x_t}$ and $u_i^D \in U_i^D$ in the subgame Γ^{x_t} and the factor-game Γ^D [12, 30].

To clarify the notation we consider the following bicriteria game in extensive form.

Example 1. A 3-player Bicriteria Multistage Game.

Let $n = 3$, $r = 2$, $P_1 = \{x_0\}$, $P_2 = \{x_1\}$, $P_3 = \{x_2\}$, $P_4 = \{x_3, x_4, x_5, x_6\}$, and

$$h(x_t) = \begin{pmatrix} h_{1/1}(x_t) & h_{2/1}(x_t) & h_{3/1}(x_t) \\ h_{1/2}(x_t) & h_{2/2}(x_t) & h_{3/2}(x_t) \end{pmatrix}, \ t \neq 0.$$

The game tree K is presented in Fig. 1. Note that the payoff $h(x_t)$ at node x_t, $t \neq 0$, is given by a matrix whose rows correspond to the criteria, and the columns correspond to the players.

Consider the following strategies $u_i(\cdot)$ in Γ^{x_0}:

$$u_1(x_0) = x_1, \ u_2(x_1) = x_2, \ u_3(x_2) = x_3.$$

Strategy profile $u = (u_1, u_2, u_3)$ generates the path $\omega(u) = (x_0, x_1, x_2, x_3)$ and the corresponding collection of the players' vector payoffs

$$H_1(u) = H_2(u) = \begin{pmatrix} 12 \\ 12 \end{pmatrix},$$

$$H_3(u) = \begin{pmatrix} 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 8 \\ 3 \end{pmatrix} = \begin{pmatrix} 12 \\ 10 \end{pmatrix}.$$

There are two subgames Γ^{x_1} and Γ^{x_2} along the trajectory $\omega(u)$. The profile of strategies $u_2^{x_1}(x_1) = x_2$, $u_3^{x_1}(x_2) = x_3$ generates the path $\omega^{x_1}(u^{x_1}) = (x_1, x_2, x_3)$ in the subgame Γ^{x_1} while $\underline{\omega}^{x_1}(u) = (x_0, x_1)$. The formulae (2) and (3) for player 3 take the form

$$H_3^{x_1}(u^{x_1}) = \tilde{h}_3^{x_1}(\omega^{x_1}(u^{x_1})) = \sum_{\tau=2}^{3} h_3(x_\tau(u^{x_1})) =$$

$$= \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 8 \\ 3 \end{pmatrix} = \begin{pmatrix} 10 \\ 6 \end{pmatrix},$$

$$H_3(u) = \tilde{h}_3(\underline{\omega}^{x_1}(u)) + \tilde{h}_3^{x_1}(\omega^{x_1}(u^{x_1})) =$$
$$= \begin{pmatrix} 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 10 \\ 6 \end{pmatrix} = \begin{pmatrix} 12 \\ 10 \end{pmatrix}.$$

We will return to this example in Section 3 when examining the strong time consistency property.

Let $a, b \in R^m$ and $k = 1, \dots, m$. To compare the vectors we use the following notations for vector preferences:

$$\begin{cases} a \geqq b, & \text{if } a_k \geqslant b_k, \forall k = 1, ..., m, \\ a > b, & \text{if } a_k > b_k, \forall k = 1, ..., m, \\ a \geq b, & \text{if } a \geqq b \text{ and } a \neq b. \end{cases}$$

The last inequality implies that the vector b is Pareto dominated by vector a (and hence b is called "inefficient").

When the players agree to cooperate in the game Γ^{x_0}, first they have to maximize the total payoffs vector $\sum_{i=1}^n H_i(u)$ w.r.t. the binary relation $\geq$. Let $PO(\Gamma^{x_0})$ be the set of all Pareto efficient pure strategy profiles from U, i.e.:

$$u \in PO(\Gamma^{x_0}) \; if \; \nexists \; v \in U : \sum_{i \in N} H_i(v) \geq \sum_{i \in N} H_i(u). \quad (4)$$

In general, the nonempty set $PO(\Gamma^{x_0})$ consists of multiple strategy profiles (see, e.g., [35, 30]), and the problem of choosing a particular Pareto efficient strategy profile arises.

3. Pareto optimal solution

3.1. Strong time consistency property

In [15], we proved that the set $PO(\Gamma^{x_0})$ of all Pareto efficient pure strategy profiles for a multistage multicriteria game Γ^{x_0} satisfies the time consistency property. This implies that if the players choose certain Pareto optimal strategy profile u that generates the trajectory $\omega(u) = (x_0, x_1(u), \dots, x_t(u), x_{t+1}(u), \dots, x_T(u))$, then in any subgame Γ^{x_t} evolving along the optimal trajectory $\omega(u)$, the restrictions of the original strategies u_i to the subgame form a Pareto efficient strategy profile in this subgame.

This property can be strengthened in order to allow for deviations from the chosen optimal strategy. This extension is referred to as *strong time consistency.*

Definition 1. [30] The set $PO(\Gamma^{x_0})$ is called *strongly time consistent* if $\forall u \in PO(\Gamma^{x_0})$, $\forall \Gamma^{x_t}$, $x_t \in \omega(u)$, and for every Pareto efficient strategy profile $W^{x_t} \in PO(\Gamma^{x_t})$ in the subgame the following inclusion holds:

$$(u^D, W^{x_t}) = ((u_1^D, W_1^{x_t}), \dots, (u_n^D, W_n^{x_t})) \in PO(\Gamma^{x_0}). \quad (5)$$

This property ensures the compatibility of locally optimal behavior (W^{x_t}) in the subgame with initial optimality requirements [28, 30, 29, 13, 14]. That is, at any intermediate state x_t the players can "switch" to another strategy profile W^{x_t} that is Pareto efficient in the current subgame Γ^{x_t}. The STC guarantees that the compound behavior obtained as a result of such switching still satisfies Pareto efficiency in the original game.

However, as the following example demonstrates, a Pareto optimal strategy profile may not necessarily be strongly time consistent.

Example 1 (Continued).

Consider the following strategies in Γ^{x_0}:

$$u_1(x_0) = x_1, \; u_2(x_1) = x_2, \; u_3(x_2) = x_3;$$
$$\sum_{i \in N} H_i(u) = \begin{pmatrix} 36 \\ 34 \end{pmatrix};$$

$$v_1(x_0) = x_4, \; v_2(x_1) = x_2, \; v_3(x_2) = x_3;$$
$$\sum_{i \in N} H_i(v) = \begin{pmatrix} 32 \\ 39 \end{pmatrix}.$$

Strategy profile $u \in PO(\Gamma^{x_0})$ generates the trajectory $\omega(u) = (x_0, x_1, x_2, x_3)$. The subgame Γ^{x_1} has two Pareto optimal strategy profiles u^{x_1} and $W^{x_1}(W_2^{x_1}(x_1) = x_2, W_3^{x_1}(x_2) = x_6)$:

$$\sum_{i \in N} H_i^{x_1}(u^{x_1}) = \begin{pmatrix} 30 \\ 18 \end{pmatrix}, \; \sum_{i \in N} H_i^{x_1}(W^{x_1}) = \begin{pmatrix} 26 \\ 21 \end{pmatrix}.$$

The compound strategy profile (u^D, W^{x_1}) in the original game Γ^{x_0} generates the trajectory

$$\lambda(u^D, W^{x_1}) = (x_0, x_1(u), x_2(W^{x_1}), x_6(W^{x_1})) =$$
$$= (x_0, x_1, x_2, x_6),$$

and the corresponding total payoff vector

$$\sum_{i \in N} H_i(u^D, W^{x_1}) = \begin{pmatrix} 32 \\ 37 \end{pmatrix}$$

which is not Pareto optimal in Γ^{x_0}. Hence, the following proposition holds.

Proposition 1. The set $PO(\Gamma^{x_0})$ in multicriteria multistage game Γ^{x_0} does not satisfy the strong time consistency property.

3.2. The Minimal Sum of Relative Deviations rule

To make a precise prediction on the players behavior in multicriteria game one need to specify a rule γ which all the players should use in order to select the time consistent cooperative path $\bar{\omega} = \omega(\bar{u})$ and the corresponding Pareto efficient strategy profile $\bar{u} \in PO(\Gamma^{x_0})$. Now we are ready to formally introduce such a specific rule which is applicable for a wide class of multicriteria games with positive payoffs. We will refer to this rule as the *minimal sum of relative deviations (MSRD)* rule.

Denote by $H_{N|k}(u) = \sum\limits_{i \in N} H_{i|k}(u)$ the sum of all players' payoffs w.r.t. the criterion k, $h_{N|k}(x_\tau) = \sum\limits_{i \in N} h_{i|k}(x_\tau)$,

$x_\tau \in P \setminus \{x_0\}$. Let $H_k^* = \max_{u\in U} H_{N|k}(u)$. The vector $(H_1^*, \dots, H_r^*)$ can be interpreted as the vector of ideal payoffs for the grand coalition N (see, e.g., [23, 35, 30, 43]).

Definition 2. According to the *MSRD rule* the players have to select a Pareto efficient pure strategy profile $\overline{u}$ which minimizes the sum of relative deviations w.r.t. each criterion from ideal payoffs vector H^*. Namely,

$$\min_{v\in U}\sum_{k=1}^{r}\frac{H_k^*-H_{N|k}(v)}{H_k^*}=\sum_{k=1}^{r}\frac{H_k^*-H_{N|k}(u)}{H_k^*},$$

or

$$\begin{aligned} u \in \arg\max_{v\in U}\sum_{k=1}^{r}\frac{1}{H_k^*}\cdot H_{N|k}(v) = \\ = \arg\max_{v\in U}\sum_{k=1}^{r}\mu_k\cdot H_{N|k}(v), \end{aligned} \tag{6}$$

where $\mu_k = \dfrac{1}{H_k^*} > 0,\ \ k = 1, \dots, r.$

Denote by $\overline{PO}(\Gamma^{x_0})$ the nonempty set of strategy profiles $u \in U$ which satisfy (6). If $|\{\omega(u),\ u \in \overline{PO}(\Gamma^{x_0})\}| = 1$, let the players choose any strategy profile $\overline{u} \in \overline{PO}(\Gamma^{x_0})$, since any such strategy profile generates the same cooperative path $\bar\omega = \omega(\bar u) = (\bar x_0, \bar x_1, \dots, \bar x_T)$, where $\bar x_0 = x_0$.

Otherwise note that every trajectory $\omega^l \in \{\omega(u), u \in \overline{PO}(\Gamma^{x_0})\}$, $l = 1, \dots, \bar m$; $\bar m > 1$ leads to some terminal node $x_{T_l} = z^j \in P_{n+1}$. Let the players choose terminal node $x_{T_l} = z^j$ with minimal number j, the corresponding path $\bar\omega = \omega(\bar u) = (\bar x_0, \dots, z^j)$ from $\{\omega(u),\ u \in \overline{PO}(\Gamma^{x_0})\}$ with terminal node z^j be the *cooperative trajectory (path)*, and the strategy profile $\bar u \in \overline{PO}(\Gamma^{x_0})$ such that $\omega(\bar u) = \bar\omega$ be the *optimal cooperative strategy profile.*

Remark 1. Strictly speaking, the MSRD rule defined above determines uniquely only the optimal cooperative trajectory $\bar\omega$ but not the optimal strategy profile. The reason for this is that in the definition of a pure strategy for extensive game there is a certain redundancy (since the player i's behavior at some nodes may not affect the outcome of the game, see [12, 24, 30] for details). However the strategy profile $\bar u$ chosen by the MSRD rule uniquely determines the players' behaviour at all nodes $\bar x_t \in \bar\omega$.

Remark 2. One can readily check that $u \in PO(\Gamma^{x_0})$ if u satisfies (6).

We will suppose in this paper that all the players have agreed to employ the minimal sum of relative deviations rule in order to find the cooperative path $\bar\omega = \omega(\bar u) = (\bar x_0, \dots, \bar x_T)$ generated by the cooperative strategy profile $\bar u \in PO(\Gamma^{x_0})$.

Denote by

$$\operatorname*{Max}^{\mu}_{u\in U}\sum_{i\in N}H_i(u)=\sum_{i\in N}H_i(\bar u) \tag{7}$$

the maximal (namely in the sense of minimal sum of relative deviations rule) total vector payoff. In addition, suppose that at every subgame $\Gamma^{\bar x_t}$, $\bar x_t \in \bar\omega$, players choose the strategy profile $u^{\bar x_t} \in U^{\bar x_t}$ such that

$$u^{\bar x_t} \in \arg\max_{v^{\bar x_t}\in U^{\bar x_t}}\sum_{k=1}^{r}\mu_k\cdot H_{N|k}^{\bar x_t}(v^{\bar x_t}), \tag{8}$$

where the coefficients $\mu_k = \dfrac{1}{H_k^*}$ are the same as in (6).

We will write $\overline{PO}(\Gamma^{\bar x_t})$ to denote the set of all strategy profiles $u^{\bar x_t} \in U^{\bar x_t}$ that satisfy (8). If strategy profiles from $\overline{PO}(\Gamma^{\bar x_t})$ generate different trajectories in $\Gamma^{\bar x_t}$, i.e., $|\{\omega^{\bar x_t}(u^{\bar x_t}), u^{\bar x_t} \in \overline{PO}(\Gamma^{\bar x_t})\}| \geq 2$, the players choose the cooperative path $\bar\omega^{\bar x_t}$ in the subgame using the same approach as in the original game Γ^{x_0} (minimal number j of the terminal node z^j).

Proposition 2. The particular Pareto efficient solution based on the MSRD rule is time consistent. Namely, suppose that $\bar u \in U$ satisfies (6), and $\bar\omega = \omega(\bar u) = (\bar x_0, \dots, \bar x_T)$ is the cooperative path in Γ^{x_0}. Then for each subgame $\Gamma^{\bar x_t}$, $\bar x_t \in \bar\omega$ along the optimal cooperative path, it holds that

$$\bar u^{\bar x_t} = (\bar u_1^{\bar x_t}, \dots, \bar u_n^{\bar x_t}) \in \arg\max_{u^{\bar x_t}\in U^{\bar x_t}}\sum_{k=1}^{r}\mu_k\cdot H_{N|k}^{\bar x_t}(u^{\bar x_t}), \tag{9}$$

while $\bar\omega^{\bar x_t} = (\bar x_t, \bar x_{t+1}, \dots, \bar x_T)$ is the corresponding cooperative path in the subgame $\Gamma^{\bar x_t}$.

Proof. Suppose that (9) does not hold, i.e., there exists $u^{\bar x_t} \in U^{\bar x_t}$ such that

$$\sum_{k=1}^{r}\mu_k\cdot H_{N|k}^{\bar x_t}(\bar u^{\bar x_t}) < \sum_{k=1}^{r}\mu_k\cdot H_{N|k}^{\bar x_t}(u^{\bar x_t}). \tag{10}$$

Let $\lambda^{\bar x_t} = (\bar x_t, x_{t+1}(u^{\bar x_t}), \dots, x_\Lambda(u^{\bar x_t}))$ be the trajectory in the subgame $\Gamma^{\bar x_t}$ generated by $u^{\bar x_t}$. Then (10) takes the form

$$\sum_{k=1}^{r}\mu_k\cdot\sum_{\tau=t+1}^{T}h_{N|k}(\bar x_\tau) < \sum_{k=1}^{r}\mu_k\cdot\sum_{\tau=t+1}^{\Lambda}h_{N|k}(x_\tau(u^{\bar x_t})). \tag{11}$$

Let, furthermore, $u_i = (\bar u_i^D, u_i^{\bar x_t})$ denote the "compound" strategy in Γ^{x_0}. Then $u = (u_1, \dots, u_n)$ generates the trajectory $\lambda = \underline{\bar\omega}^{\bar x_t} \cup \lambda^{\bar x_t} = (\bar x_0, \dots, \bar x_t, x_{t+1}(u^{\bar x_t}), \dots, x_\Lambda(u^{\bar x_t}))$ in Γ^{x_0}.

Adding $\sum\limits_{k=1}^{r}\mu_k\sum\limits_{\tau=1}^{t}h_{N|k}(\bar x_\tau)$ to both sides of (11) and taking into account (3) we get

$$\sum_{k=1}^{r}\mu_k\cdot H_{N|k}(\bar u) < \sum_{k=1}^{r}\mu_k\cdot H_{N|k}(u)$$

for some $u \in U$. The last inequalitiy contradicts the condition $\bar u \in \overline{PO}(\Gamma^{x_0})$, hence (9) is valid.

Arguing in a similar way (for the case when $|\{\omega^{\bar x_t}(u^{\bar x_t}), u^{\bar x_t} \in \overline{PO}(\Gamma^{\bar x_t})\}| > 1$) one can verify that $\bar\omega^{\bar x_t} = (\bar x_t, \dots, \bar x_T)$ — a fragment of the optimal path $\bar w$, starting at $\bar x_t$ — remains the optimal cooperative path in the subgame $\Gamma^{\bar x_t}$. □

4. Constructing a characteristic function for multicriteria game

To design a vector-valued characteristic function for multistage multicriteria game Γ^{x_0} we employ the method described in [7]. The corresponding ζ-characteristic function is relatively friendly computable and is proved to satisfy the superadditivity property for (single-criterion) cooperative differential n-person games. This approach implies the following two-stage scheme: first the players select the optimal cooperative strategy profile $\bar{u} \in PO(\Gamma^{x_0})$ using MSRD rule. Then, we assume that all players from S use the optimal cooperative strategies $\bar{u}_j$, while the other players (from $N \backslash S$) seek to minimize (in the sense of MSRD rule) the total payoffs vector of the players from coalition S:

$$V(S) = \begin{cases} \bar{0} \in R^r, & S = \emptyset, \\ \operatorname*{Min}^{\mu}_{u_j, j \in N\backslash S} \sum\limits_{i \in S} H_i(\bar{u}_S, u_{N\backslash S}), & S \subset N, \\ \operatorname*{Max}^{\mu}_{u \in U} \sum\limits_{i \in N} H_i(u), & S = N. \end{cases} \tag{12}$$

Namely, the players from $N \setminus S$ solve the following optimization problem:

$$\min_{u_j, j \in N\backslash S} \sum_{k=1}^{r} \mu_k \cdot \sum_{i \in S} H_{i|k}(\bar{u}_S, u_{N\backslash S}) = \\ = \sum_{k=1}^{r} \mu_k \cdot \sum_{i \in S} H_{i|k}(\bar{u}_S, \underline{u}_{N\backslash S}). \tag{13}$$

It is worth noting that coefficients μ_k in (13) are the same as in (6).

Again, if for all $\underline{u}_{N\backslash S}$ satisfying (13) $|\omega(\bar{u}_S, \underline{u}_{N\backslash S})| = 1$, the players from $N \setminus S$ can choose any $\underline{u}_{N\backslash S}$ meeting (13). Otherwise, they are expected to choose the path $\omega(\bar{u}_S, \underline{u}_{N\backslash S})$ whose terminal node z^j has minimal number j, and corresponding bundle of strategies $(\underline{u}_j)_{j \in N\backslash S}$. Let

$$\operatorname*{Min}^{\mu}_{u_j, j \in N\backslash S} \sum_{i \in S} H_i(\bar{u}_S, u_{N\backslash S}) = \sum_{i \in S} H_i(\bar{u}_S, \underline{u}_{N\backslash S})$$

denote the minimal (in the sense of the MSRD rule) total vector payoff of the coalition S.

Definition 3. The vector-valued characteristic function $V(S)$ in multicriteria game Γ satisfies *weak superadditivity property* if for all $S_1, S_2 \subseteq N$ with $S_1 \cap S_2 = \emptyset$ the following vector inequality

$$V(S_1 \cup S_2) \leq V(S_1) + V(S_2) \tag{14}$$

does not hold.

Proposition 3. The vector-valued characteristic function (12) is weakly superadditive.

Proof. Consider proper coalitions $S_1, S_2 \subset N$, $S_1 \cap S_2 = \emptyset$, $S = S_1 \cup S_2 \subset N$. Let

$$\operatorname*{Min}^{\mu}_{u_j, j \in N\backslash S_1} \sum_{i \in S_1} H_i(\bar{u}_{S_1}, u_{N\backslash S_1}) = \sum_{i \in S_1} H_i(\bar{u}_{S_1}, u'_{N\backslash S_1}),$$

$$\operatorname*{Min}^{\mu}_{u_j, j \in N\backslash S_2} \sum_{i \in S_2} H_i(\bar{u}_{S_2}, u_{N\backslash S_2}) = \sum_{i \in S_2} H_i(\bar{u}_{S_2}, u''_{N\backslash S_2}).$$

Then for all possible $u_{N\backslash S_1} \subset U_{N\backslash S_1}$ and $u_{N\backslash S_2} \subset U_{N\backslash S_2}$

$$\begin{cases} \sum\limits_{k=1}^{r} \mu_k \sum\limits_{i \in S_1} H_{i|k}(\bar{u}_{S_1}, u'_{N\backslash S_1}) \leq \sum\limits_{k=1}^{r} \mu_k \sum\limits_{i \in S_1} H_{i|k}(\bar{u}_{S_1}, u_{N\backslash S_1}), \\ \sum\limits_{k=1}^{r} \mu_k \sum\limits_{i \in S_2} H_{i|k}(\bar{u}_{S_2}, u''_{N\backslash S_2}) \leq \sum\limits_{k=1}^{r} \mu_k \sum\limits_{i \in S_2} H_{i|k}(\bar{u}_{S_2}, u_{N\backslash S_2}). \end{cases} \tag{15}$$

Suppose that inequality (14) is valid, i.e.

$$\sum_{i \in S} H_{i|k}(\bar{u}_S, \underline{u}_{N\backslash S}) \leq \sum_{i \in S_1} H_{i|k}(\bar{u}_{S_1}, u'_{N\backslash S_1}) + \\ + \sum_{i \in S_2} H_{i|k}(\bar{u}_{S_2}, u''_{N\backslash S_2}), k = 1, \ldots, r; \tag{16}$$

and at least one inequality from (16) is strict.

If we multiply each inequality from (16) by $\mu_k > 0$ and then sum up all the inequalities, we get

$$\sum_{k=1}^{r} \mu_k \sum_{i \in S} H_{i|k}(\bar{u}_S, \underline{u}_{N\backslash S}) = \sum_{k=1}^{r} \mu_k \sum_{i \in S_1} H_{i|k}(\bar{u}_{S_1}, \bar{u}_{S_2}, \underline{u}_{N\backslash S}) + \\ + \sum_{k=1}^{r} \mu_k \sum_{i \in S_2} H_{i|k}(\bar{u}_{S_1}, \bar{u}_{S_2}, \underline{u}_{N\backslash S}) < \\ < \sum_{k=1}^{r} \mu_k \sum_{i \in S_1} H_{i|k}(\bar{u}_{S_1}, u'_{N\backslash S_1}) + \sum_{k=1}^{r} \mu_k \sum_{i \in S_2} H_{i|k}(\bar{u}_{S_2}, u''_{N\backslash S_2}).$$

The last inequality contradicts system (15).

Hence, (14) does not hold, and $V(S)$ is weakly superadditive.

Obviously, for the coalition $S = \emptyset$ and $S = N$ the weak superadditivity property holds trivially. □

5. Long-term implementation of the cooperative solution

5.1. The imputation distribution procedure and its properties

An imputation distribution procedure (or payment schedule) is known to be a useful method to implement a long-term cooperative agreement in a dynamic game (see, e.g., [28, 31, 30, 37, 17, 18, 8]). Below, we introduce a number of properties that can be used to characterize an imputation distribution procedure. For certainty, we will formulate these properties for the Shapley value and later on for the core, although they can be readily generalized for an arbitrary imputation (see, for instance, [23, 24, 30, 17]) from a cooperative solution.

The core and the Shapley value were extended to cooperative multicriteria games in [1, 4, 33] and [34]. Denote by $\Gamma^{x_0}(N, V^{x_0})$ a multicriteria game Γ^{x_0} with the vector-valued characteristic function V^{x_0} defined by (12).

Definition 4. [40] The *Shapley value* of $\Gamma^{x_0}(N, V^{x_0})$ is a vector φ^{x_0} which is defined for every player $i \in N$ as

$$\varphi_i^{x_0} = \sum_{S \subset N, i \in S} \frac{(n - |S|)!(|S| - 1)!}{n!} (V^{x_0}(S) - V^{x_0}(S \setminus \{i\})). \tag{17}$$

Note that the Shapley value for a cooperative multicriteria game is proved to satisfy the so-called efficiency property [40, 34], i.e.:

$$\sum_{i=1}^{n} \varphi_i^{x_0} = V^{x_0}(N) = \sum_{\tau=1}^{T}\sum_{i=1}^{n} h_i(\bar{x}_\tau). \tag{18}$$

However, if a vector-valued characteristic function for multicriteria game does not satisfy the strong (component-wise) superadditivity property, the Shapley value may not necessarily satisfy the so-called individual rationality constraint: $\varphi_i^{x_0} \geqq V^{x_0}(\{i\})$.

For every $\bar{x}_t \in \bar{\omega}(\bar{u})$, denote by $\Gamma^{\bar{x}_t}(N, V^{\bar{x}_t})$, $t = 0, \ldots, T-1$, a subgame along the cooperative path with the subgame characteristic function $V^{\bar{x}_t}$ which can be computed for this subgame using the same approach as in (12). Note that

$$V^{\bar{x}_t}(N) = \sum_{\tau=t+1}^{T}\sum_{i\in N} h_i(\bar{x}_\tau). \tag{19}$$

Let $(\varphi_i^{\bar{x}_t})_{i\in N}$ denote the Shapley value for the subgame $\Gamma^{\bar{x}_t}(N, V^{\bar{x}_\tau})$. Denote by $\beta = \{\beta_{i/k}(\bar{x}_\tau)\}, i = 1, \ldots, n$; $k = 1, \ldots, r$; $\tau = 1, \ldots, T$ the IDP (or the payment schedule) [28, 31, 30, 17, 37]. The payment schedule based approach means that all the players have agreed to allocate (according to some specific rule called IDP) the total cooperative vector payoff $\sum_{i\in N} H_i(\bar{u}) = V^{x_0}(N)$ between the players along the cooperative path $\bar{\omega}(\bar{u})$. Then $\beta_{i/k}(\bar{x}_\tau)$ denotes the actual current payment that the i-th player receives at node $\bar{x}_\tau$ w.r.t. criterion k (instead of $h_{i/k}(\bar{x}_\tau)$) when the players apply the imputation distribution procedure β.

If at the initial time each player is satisfied with the respective share $\varphi_i^{x_0}$ of the total payoff $V^{x_0}(N)$, then an appropriate payment schedule can be used to keep the player interested in cooperation at any intermediate time, i.e., in any subgame Γ^{x_t}.

Definition 5. [17] The imputation distribution procedure $\beta = \{\beta_{i/k}(\overline{x}_t)\}$ satisfies the *efficiency condition* if

$$\sum_{t=1}^{T} \beta_i(\bar{x}_t) = \varphi_i^{x_0}, i = 1, \ldots, n. \tag{20}$$

Equation (20) means that the sum of the player i's actual current payments along the optimal cooperative path equals to the total amount to be obtained by the player i according to the cooperative solution. Then the payment schedule for every player can be reasonably interpreted as a rule for the step-by-step allocation of the player i's optimal payoff.

Definition 6. The imputation distribution procedure β satisfies the *strict balance condition* if $\forall t = 1, .., T; \forall k = 1, ..., r$

$$\sum_{\tau=1}^{t}\sum_{i=1}^{n} \beta_{i/k}(\bar{x}_\tau) = \sum_{\tau=1}^{t}\sum_{i=1}^{n} h_{i/k}(\bar{x}_\tau). \tag{21}$$

This condition ensures that at any intermediate state $\bar{x}_t$ along the optimal trajectory the players have collected exactly the amount of payments that is needed to implement the procedure β.

The next useful dynamic property of a payment schedule — the IBP condition, introduced in [45] — was extended to multicriteria cooperative games in [17, 15] .

Definition 7. [17] The imputation distribution procedure $\beta = \{\beta_{i/k}(\bar{x}_t)\}$, $i = 1, \ldots, n$; $t = 1, \ldots, T$ in $\Gamma^{x_0}(N, V^{x_0})$ satisfies the *strong (or component-wise) irrational-behavior-proof (IBP) condition for individual players*, if for all $i \in N$ and for any $t = 1, \ldots, T-1$ it holds that:

$$\sum_{\tau=1}^{t} \beta_i(\bar{x}_\tau) + V^{\bar{x}_t}(\{i\}) \geqq V^{x_0}(\{i\}). \tag{22}$$

Vector inequality (22) implies that every player has a reasonable incentive to cooperate (at least until the intermediate node $\bar{x}_t$ will be reached) even if he anticipates that the cooperation can be destroyed at the node $\bar{x}_t$ because of the "irrational behavior" of some other players.

Definition 8. [15] The imputation distribution procedure β satisfies the *strong irrational-behavior-proof (IBP) condition (for coalitions)* if $\forall\ S \subset N$, $|S| > 1$, for any $t = 1, \ldots, T-1$ it holds that:

$$\sum_{i\in S}\sum_{\tau=1}^{t} \beta_i(\bar{x}_\tau) + V^{\bar{x}_t}(S) \geqq V^{x_0}(S). \tag{23}$$

Two other advantageous properties of payment schedules in multicriteria game – time consistency and non-negativity – were investigated in [17, 18].

Definition 9. [17] The imputation distribution procedure $\beta = \{\beta_{i/k}(\bar{x}_t)\}$ satisfies the *time consistency* property if

$$\sum_{\tau=1}^{t} \beta_i(\bar{x}_\tau) + \varphi_i^{\bar{x}_t} = \varphi_i^{x_0} \quad \forall t = 1, \ldots, T-1,\ i \in N.$$

This constraint means that the total payment received by the i-th player along the cooperative path before entering the intermediate node $\bar{x}_t$ plus the i-th component of the subgame Shapley value is equal to the player i's optimal payoff to be obtained in the original game.

Definition 10. The imputation distribution procedure β satisfies the *non-negativity constraint* if

$$\beta_{i/k}(\bar{x}_t) \geqslant 0,\ i = 1, \ldots, n;\ k = 1, \ldots, r;\ t = 1, \ldots, T.$$

We note that there could be different payment schedules that may or may not satisfy the properties listed above. Below we consider three payment schedules and discuss their properties.

The first one, called the *incremental* imputation distribution procedure is formulated as follows, [28, 17]:

$$\begin{aligned} &\beta_i(\bar{x}_t) = \varphi_i^{\bar{x}_{t-1}} - \varphi_i^{\bar{x}_t},\ t = 1, \ldots, T-1; \\ &\beta_i(\bar{x}_T) = \varphi_i^{\bar{x}_{T-1}}. \end{aligned} \tag{24}$$

This payment schedule was studied extensively for different classes of single-criterion dynamic games (see for instance [30]). The incremental payment schedule is designed to satisfy time consistency, but in general this IDP does not satisfy the non-negativity condition (two approaches how to overcome this disadvantage were described in [17, 8]).

The *proportional* imputation distribution procedure is defined as

$$\beta_{i/k}(\bar{x}_t) = \frac{\sum_{i=1}^{n} h_{i/k}(\bar{x}_t)}{V_k^{x_0}(N)} \cdot \varphi_{i/k}^{\bar{x}_0}, t = 1, \dots, T. \tag{25}$$

This payment schedule implies the proportional allocation of the total current vector payoff at every node along the optimal cooperative trajectory. The proportional IDP obviously satisfies the non-negativity constraint for multicriteria games with positive payoffs but does not satisfy time consistency (see, e.g., [28, 17]).

Finally, we consider a novel payment schedule as described below.

Definition 11. The *V-incremental* payment schedule is defined as follows

$$\begin{aligned} \beta_i(\bar{x}_t) &= V^{\bar{x}_{t-1}}(\{i\}) - V^{\bar{x}_t}(\{i\}) + \\ &+\frac{1}{n}\left(V^{\bar{x}_{t-1}}(N) - \sum_{i\in N} V^{\bar{x}_{t-1}}(\{i\})\right) - \\ &-\frac{1}{n}\left(V^{\bar{x}_t}(N) - \sum_{i\in N} V^{\bar{x}_t}(\{i\})\right), t = 1, \dots, T-1; \\ \beta_i(\bar{x}_T) &= \varphi_i^{x_0} - \sum_{\tau=1}^{T-1} \beta_i(\bar{x}_\tau). \end{aligned} \tag{26}$$

This payment schedule is neither non-negative nor time consistent in general, but it is the most flexible one as the players may postpone the choice of a particular supporting imputation from the core up to the final stage of a game.

To conclude, we note that all these payment schedules always satisfy the efficiency (20) and the strict balance conditions (21).

5.2. Choice of supporting imputation for the core

In the following we will consider the core as the multiple cooperative solution.

Definition 12. [33] The *core* $C(V^{x_0})$ of the cooperative game $\Gamma^{x_0}(N, V^{x_0})$ is a set of all imputations $\alpha = (\alpha_1, \dots, \alpha_n)$ satisfying the inequalities

$$\sum_{i\in S} \alpha_i = \alpha_S \geqq V^{x_0}(S), \ \forall S \subset N. \tag{27}$$

Likewise, for every subgame $\Gamma^{\bar{x}_t}(N, V^{\bar{x}_t})$, $t = 1, \dots, T-1$ we denote by $C(V^{\bar{x}_t})$ the core of this subgame.

Let $A[r \times n]$ denote the set of all matrices with real components which contain r lines and n columns, $a \in A[r \times n]$, $B \subset A[r \times n]$. We employ the notation $a \oplus B = \{a + b : b \in B\}$ to define the strong time consistency of the whole core [6, 15].

Definition 13. [6] The core $C(V^{x_0})$ satisfies the *strong time consistency* property in the game $\Gamma^{x_0}(N, V^{x_0})$ if

1. $C(V^{\bar{x}_t}) \neq \emptyset$ for all $t = 0, \dots, T-1$.
2. There exists an imputation $\bar{\alpha} \in C(V^{x_0})$ and imputation distribution procedure $\beta = \{\beta_{i/k}(\bar{x}_t)\}$, $t = 1, \dots, T$, such that $\sum_{t=1}^{T} \beta_i(\bar{x}_t) = \bar{\alpha}_i$, $i = 1, \dots, n$, and

$$C(V^{x_0}) \supset \sum_{\tau=1}^{t} \beta(\bar{x}_\tau) \oplus C(V^{\bar{x}_t}), \ t = 1, \dots, T-1. \tag{28}$$

Definition 14. The vector $\bar{\alpha}$ from the core $C(V^{x_0})$ and the imputation distribution procedure $\beta = \{\beta_{i/k}(\bar{x}_t)\}$ which provide condition (28) are called the *supporting imputation* and the *supporting IDP*, respectively.

If the core $C(V^{x_0})$ satisfies the STC it is possible to find the supporting imputation $\bar{\alpha}$ inside the core and to redistribute it over time using payment schedule β such that any deviation from this supporting solution to any other imputation from the subgame core $C(V^{\bar{x}_t})$ will result to the vector of payoffs that is also contained in the core $C(V^{x_0})$.

Proposition 4. Let for every subgame $\Gamma^{\bar{x}_t}(N, V^{\bar{x}_t})$, $t = 0, \dots, T-1$ formed along the cooperative path $\bar{\omega}$ the core $C(V^{\bar{x}_t})$ be non-empty and contain the Shapley value $\varphi^{\bar{x}_t}$. Let a payment schedule $\beta = \{\beta_{i/k}(\bar{x}_t)\}$ for Shapley value satisfy the efficiency condition (20), the strict balance condition (21), and the strong IBP conditions for individual players (22) and for coalitions(23).

Then the whole core $C(V^{x_0})$ in the original game $\Gamma(N, V^{x_0})$ satisfies the STC while the Shapley value $\varphi_i^{x_0}$ and payment schedule $\beta = \{\beta_{i/k}(\bar{x}_t)\}$ are used as a supporting imputation $\bar{\alpha}$ and supporting imputation distribution procedure, respectively.

Proof. To verify (28) let us select an arbitrary $t \in \{1, \dots, T-1\}$ and any imputation $\alpha^t \in C(V^{\bar{x}_t})$ from the subgame $\Gamma^{\bar{x}_t}(N, V^{\bar{x}_t})$ core . Denote by $\hat{\alpha}_i = \sum_{\tau=1}^{t} \beta_i(\bar{x}_\tau) + \alpha_i^t$ the resulting vector of the players' payoffs in $\Gamma(N, V^{x_0})$.

We have to prove that $\hat{\alpha}$ satisfies inequalities (27), i.e. that $(\hat{\alpha}_1, \dots, \hat{\alpha}_n) \in C(V^{x_0})$.

Using the strict balance condition (21), (18) and (19) we obtain

$$\begin{aligned} \sum_{i\in N}\sum_{\tau=1}^{t} \beta_i(\bar{x}_\tau) &= \sum_{i\in N}\sum_{\tau=1}^{T} \beta_i(\bar{x}_\tau) - \sum_{i\in N}\sum_{\tau=t+1}^{T} \beta_i(\bar{x}_\tau) = \\ &= \sum_{i\in N}\sum_{\tau=1}^{T} h_i(\bar{x}_\tau) - \sum_{i\in N}\sum_{\tau=t+1}^{T} h_i(\bar{x}_\tau) = V^{x_0}(N) - V^{\bar{x}_t}(N). \end{aligned}$$

Then

$$\begin{aligned} \sum_{i\in N} \hat{\alpha}_i &= \sum_{i\in N}\sum_{\tau=1}^{t} \beta_i(\bar{x}_\tau) + \sum_{i\in N} \alpha_i^t = \\ &= V^{x_0}(N) - V^{\bar{x}_t}(N) + V^{\bar{x}_t}(N) = V^{x_0}(N). \end{aligned}$$

Since α^t is an imputation in the subgame $\Gamma^{\bar{x}_t}(N, V^{\bar{x}_t})$, and taking into account the strong IBP condition for individual players (22) we get

$$\hat{\alpha}_i = \sum_{\tau=1}^{t} \beta_i(\bar{x}_\tau) + \alpha_i^t \geqq V^{x_0}(\{i\}) - V^{\bar{x}_t}(\{i\}) + \alpha_i^t \geqq$$
$$\geqq V^{x_0}(\{i\}) - V^{\bar{x}_t}(\{i\}) + V^{\bar{x}_t}(\{i\}) \geqq V^{x_0}(\{i\}).$$

Thus, $\hat{\alpha}$ is exactly an imputation in the original game $\Gamma(N, V^{x_0})$.

Since $\alpha^t \in C(V^{\bar{x}_t})$, using the strong IBP condition for coalitions (23) we obtain

$$\sum_{i\in S} \hat{\alpha}_i = \sum_{i\in S}\sum_{\tau=1}^{t} \beta_i(\bar{x}_\tau) + \sum_{i\in S} \alpha_i^t \geqq$$
$$\geqq V^{x_0}(S) - V^{\bar{x}_t}(S) + \sum_{i\in S} \alpha_i^t \geqq$$
$$\geqq V^{x_0}(S) - V^{\bar{x}_t}(S) + V^{\bar{x}_t}(S) \geqq V^{x_0}(S),\ \forall S \subset N,\ |S| > 1.$$

Hence, the resulting vector of the players' payoffs $\hat{\alpha}$ despite of the deviation made in the subgame $\Gamma^{\bar{x}_t}(N, V^{\bar{x}_t})$ still belongs to the core $C(V^{x_0})$ of the original game. □

It is worth noting that Proposition 4 does not depend on a particular optimal imputation the players select within the core given that this imputation and corresponding payment schedule satisfy the strong IBP conditions (22) and (23).

Any of the payment schedules introduced in Sec. 5.1 can be used as a supporting imputation distribution procedure. In this case, one only need to verify whether the Shapley value belongs to the core at every subgame along the cooperative path and whether it satisfies the strong IBP conditions for individual players (22) and for coalitions (23).

6. Example. Implementation of three payment schedules for given multicriteria game

Consider 3-person game with 3 criteria: $n = 3$, $r = 3$, $P_1 = \{x_0\}$, $P_2 = \{x_1\}$, $P_3 = \{x_2\}$, $P_4 = \{x_3, x_4, x_5, x_6\}$, and

$$h(x_t) = \begin{pmatrix} h_{1/1}(x_t) & h_{2/1}(x_t) & h_{3/1}(x_t) \\ h_{1/2}(x_t) & h_{2/2}(x_t) & h_{3/2}(x_t) \\ h_{1/3}(x_t) & h_{2/3}(x_t) & h_{3/3}(x_t) \end{pmatrix}.$$

The game dynamics and payoffs are presented in Fig. 2.

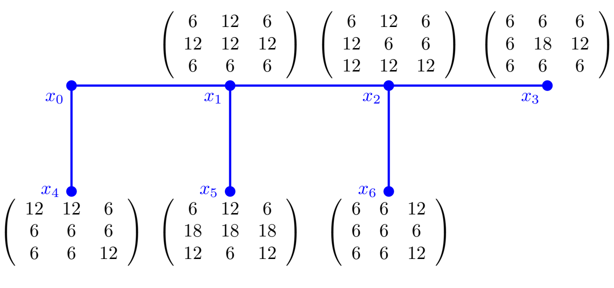

Figure 2: The 3-criteria game tree

There exist two Pareto optimal trajectories in this game. Using the MSRD rule the players choose the optimal cooperative strategy profile $\bar{u} = (\bar{u}_1, \bar{u}_2, \bar{u}_3)$: $\bar{u}_1(x_0) = x_1$, $\bar{u}_2(x_1) = x_2$, $\bar{u}_3(x_2) = x_3$; which generates the cooperative trajectory $\bar{\omega} = \omega(\bar{u}) = (x_0, x_1, x_2, x_3)=(\bar{x}_0, \bar{x}_1, \bar{x}_2, \bar{x}_3)$. The values of the vector-valued ζ − characteristic function (12) for the game Γ^{x_0} are

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{x_0}(S)$	12	12	6	48	24	18	66
	30	6	6	54	60	12	96
	18	6	12	48	36	18	72

and the Shapley value for original game Γ^{x_0} is

$$\varphi^{x_0} = \begin{pmatrix} 29 & 26 & 11 \\ 55 & 19 & 22 \\ 35 & 20 & 17 \end{pmatrix}.$$

The vector-valued ζ − characteristic functions and the respective Shapley values for the subgames along the cooperative path $\bar{\omega}$ can be constructed using the same approach (12).

The subgame $\Gamma^{x_1}(N, V^{x_1})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{x_1}(S)$	6	18	6	30	12	30	42
	18	12	18	30	36	42	60
	12	18	12	36	24	36	54

$$\varphi^{x_1} = \begin{pmatrix} 9 & 24 & 9 \\ 18 & 18 & 24 \\ 15 & 24 & 15 \end{pmatrix}.$$

The subgame $\Gamma^{x_2}(N, V^{x_2})$:

S	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	N
$V^{x_2}(S)$	6	6	6	12	12	12	18
	6	6	12	12	18	30	36
	6	6	6	12	12	12	18

$$\varphi^{x_2} = \begin{pmatrix} 6 & 6 & 6 \\ 6 & 12 & 18 \\ 6 & 6 & 6 \end{pmatrix}.$$

One can readily check that the Shapley value satisfies (27) for the original game and for both subgames along the optimal cooperative path, i.e., $\varphi^{\bar{x}_t} \in C(V^{\bar{x}_t})$, $t = 0, 1, 2$. Therefore, the core $C(V^{\bar{x}_t})$ is nonempty $\forall t = 0, 1, 2$. The simplest incremental imputation distribution procedure (24)

$$\{\beta_{i/k}(\bar{x}_1)\} = \varphi^{\bar{x}_0} - \varphi^{\bar{x}_1} = \begin{pmatrix} 20 & 2 & 2 \\ 37 & 1 & -2 \\ 20 & -4 & 2 \end{pmatrix},$$

$$\{\beta_{i/k}(\bar{x}_2)\} = \begin{pmatrix} 3 & 18 & 3 \\ 12 & 6 & 6 \\ 9 & 18 & 9 \end{pmatrix}, \{\beta_{i/k}(\bar{x}_3)\} = \begin{pmatrix} 6 & 6 & 6 \\ 6 & 12 & 18 \\ 6 & 6 & 6 \end{pmatrix}$$

and the proportional payment schedule (25)

$$\{\widetilde{\beta}_{i/k}(\bar{x}_1)\} = \begin{pmatrix} 10,(54) & 9,(45) & 4 \\ 20,625 & 7,125 & 8,25 \\ 8,75 & 5 & 4,25 \end{pmatrix},$$

$$\{\widetilde{\beta}_{i/k}(\bar{x}_2)\} = \begin{pmatrix} 10,(54) & 9,(45) & 4 \\ 13,75 & 4,75 & 5,5 \\ 17,5 & 10 & 8,5 \end{pmatrix},$$

$$\{\widetilde{\beta}_{i/k}(\bar{x}_3)\} = \begin{pmatrix} 7,(90) & 7,(09) & 3 \\ 20,625 & 7,125 & 8,25 \\ 8,75 & 5 & 4,25 \end{pmatrix}$$

satisfy the strong IBP conditions for individual players (22) and for coalitions (23) for all $t = 1, 2$. Therefore, the whole core $C(V^{x_0})$ satisfies the STC, and the Shapley value φ^{x_0} can be employed as a supporting optimal imputation while the players use either the incremental imputation distribution procedure (24) or the proportional IDP (25). The choice of one of the two payment schedules can be made taking into account which of the two properties - time consistency or non-negativity - is more important for the players.

On the other hand, the V-incremental payment schedule (26)

$$\{\widetilde{\widetilde{\beta}}_{i/k}(\bar{x}_1)\} = \begin{pmatrix} 14 & 2 & 8 \\ 26 & 8 & 2 \\ 14 & -4 & 8 \end{pmatrix},$$

$$\{\widetilde{\widetilde{\beta}}_{i/k}(\bar{x}_2)\} = \begin{pmatrix} 4 & 16 & 4 \\ 12 & 6 & 6 \\ 10 & 16 & 10 \end{pmatrix}, \{\widetilde{\widetilde{\beta}}_{i/k}(\bar{x}_3)\} = \begin{pmatrix} 11 & 8 & -1 \\ 17 & 5 & 14 \\ 11 & 8 & -1 \end{pmatrix}$$

does not satisfy the strong IBP condition for coalitions for $t = 1$ (since condition (23) is violated for coalition $S = \{1, 2\}$). Therefore, in this game, V-incremental payment schedule does not ensure the sustainability of a long-term cooperative agreement and cannot be used as a supporting IDP.

7. Conclusion

In this paper, we introduced a new rule that can be used to choose an optimal and time consistent cooperative trajectory in a multicriteria multistage game. This rule can be used for a wide class of multicriteria games with positive payoffs to construct a vector-valued characteristic function that satisfies weak superadditivity property. The players are assumed to use the payment schedule based approach to guarantee the sustainability of a cooperative agreement, i.e. the players design an appropriate allocation mechanism in order to distribute the optimal payoff of each player along the cooperative trajectory.

Furthermore, we specify the minimal set of properties a payment schedule should satisfy to guarantee the STC of the whole core. Namely, if the payment schedule designed to distribute the Shapley value satisfies strict balance condition, efficiency condition and the strong IBP conditions, given that the Shapley value belongs to the core, it can be employed as a supporting imputation distribution procedure. The STC of the core implyes that a single deviation from the Shapley value to any other imputation chosen from the subgame core still lead to the payments vector from the core of the original multistage game. We discuss three payment schedules - the incremental, proportional and V-incremental IDP - and check whether they can be used as supporting imputation distribution procedures for the given multicriteria game.

An interesting research question is to provide an axiomatic characterization (see, [42, 36]) of the proposed MSRD rule for choosing a unique Pareto efficient solution in multicriteria games.

Acknowledgements

We would like to thank the anonymous Reviewers and the Editor for their valuable comments.

The reported study was funded by RFBR under the research project 18-00-00727 (18-00-00725).

References

[1] K. Bergstresser, P. Yu, Domination structures and multicriteria problems in N-person games, Theory and Decision 8 (1977) 5–48.

[2] P. Borm, S. Tijs, J. Van Den Aarssen, Pareto equilibrium in multiobjective games, Methods Oper. Res. 60 (1990) 303–312.

[3] J. Climaco, C. Romero, F. Ruiz, Preface to the Special Issue on Multiple Criteria Decision Making: Current Challenges and Future Trends, Intl. Trans. in Op. Res. 25 (2018) 759–761, https://doi.org/10.1111/itor.12515.

[4] F.R. Fernández, M.A. Hinojosa, J. Puerto, Core solutions in vector-valued games, J. Optim. Theory Appl. 112 (2002) 331–360.

[5] M. Finus, Game Theory and International Environmental Cooperation. Edward Elgar Ed, 2001.

[6] E. Gromova, The Shapley Value as a Sustainable Cooperative Solution in Differential Games of Three Players, Recent Advances in Game Theory and Applications, Static and Dynamic Game Theory: Foundations and Applications. Springer-Verlag, (2016) 67-91.

[7] Gromova, E. V. Petrosyan L.A. On an approach to constructing a characteristic function in cooperative differential games, Automation and Remote Control, Vol. 78 (9) (2017) 1680-1692.

[8] E.V. Gromova E.V., Plekhanova T.M., On the regularization of a cooperative solution in a multistage game with random time horizon, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.08.008.

[9] M. Haddad, D. Sanders, Selection of discrete multiple criteria decision making methods in the presence of risk and uncertainty, Operations Research Perspectives, Vol. 5, (2018) 357-370, https://doi.org/10.1016/j.orp.2018.10.003 .

[10] A. Haurie, A note on nonzero-sum diferential games with bargaining solution, Journal of Optimization Theory and Applications 18 (1976) 31-39.

[11] L.Krus, P.Bronisz, Solution concepts in multicriteria cooperative games without side payments, System Modeling and Optimization, J.Dolezal et al.(eds.), IFIP, Springer, Boston (1996) 300-307.

[12] H. Kuhn, Extensive games and the problem of information, Annals of Mathematical Studies 28 (1953), 193-216.

[13] D. Kuzyutin, On the problem of the stability of solutions in extensive games, Vestnik St. Petersburg Univ. Math. 4 (22) (1995) 18-23. in Russian.

[14] D. Kuzyutin, Properties of solutions of positional games with incomplete information, Journal of Computer and Systems Sciences International 34(1) (1996) 157-162.
[15] D. Kuzyutin, E. Gromova, Ya. Pankratova, Sustainable cooperation in multicriteria multistage games, Operations Research Letters, Vol. 46, issue 6, (2018) 557–562, https://doi.org/10.1016/j.orl.2018.09.004 .
[16] D. Kuzyutin , M. Nikitina , L. Razgulyaeva, On the A-equilibria properties in multicriteria extensive games, Applied Mathematical Sciences 9(92) (2015) 4565-4573.
[17] D. Kuzyutin, M. Nikitina, Time consistent cooperative solutions for multistage games with vector payoffs, Operations Research Letters, Vol. 45, issue 3, (2017) 269 – 274.
[18] D. Kuzyutin, M. Nikitina, An irrational behavior proof condition for multistage multicriteria games, in: Consrtuctive nonsmooth analysis and related topics (dedic. to the memory of V.F.Demyanov), CNSA 2017, Proceedings. IEEE, (2017) 178-181.
[19] J. Lund, E. Hanak, et al. Comparing futures for the Sacramento-San Joaquin Delta. Berkeley, CA: Univ.of California Press, 2010.
[20] K. Madani, Game theory and water resources, Journal of Hydrology, 381 (2010), 225-238.
[21] K. Madani, J.R.Lund, A Monte-Carlo game theoretic approach for multi-criteria decision making under uncertainty, Advances in Water Resources, 34 (2011), pp. 607-616.
[22] G.A. Mendoza, H. Martins, Multi-criteria decision analysis in natural resource management: A critical review of methods and new modelling paradigms, Forest Ecology and Management, 230 (2006), 1–22.
[23] H. Moulin, Axioms of cooperative decision making, Cambridg Univ. Press, 1988.
[24] R. Myerson, Game Theory. Analysis of Conflict, Harvard University Press, 1997.
[25] Y. Pankratova, S. Tarashnina, D. Kuzyutin, Nash equilibria in a group pursuit game, Applied Mathematical Sciences 10 (17) (2016) 809-821.
[26] E. Parilina , G. Zaccour, Approximated cooperative equilibria for games played over event trees, Operations Research Letters 43 (5) 2015 507-513.
[27] L. Petrosyan, Stable solutions of differential games with many participants, Vestnik of Leningrad University 19 (1977) 46–52. (in Russian).
[28] L. Petrosyan, N. Danilov, Stability of the solutions in nonantagonistic differential games with transferable payoffs, Vestnik of Leningrad University 1 (1979) 52-59. (in Russian).
[29] Petrosyan L.A., Kuzyutin D.V., On the stability of E-equilibrium in the class of mixed strategies, Vestnik St.Petersburg Univ. Math. 3 (15) (1995) 54-58. in Russian.
[30] L. Petrosyan, D. Kuzyutin, Consistent solutions of positional games, Saint Petersburg University Press, 2008.(in Russian).
[31] L. Petrosyan, G. Zaccour, Time-consistent Shapley value allocation of pollution cost reduction, Journal of Economic Dynamics and Control 27(3) (2003) 381-398.
[32] O. L. Petrosyan, E.V. Gromova, S.V. Pogozhev, Strong time-consistent subset of the core in cooperative differential games with finite time horizon, Automation and Remote Control, 79(10) (2018) 1912-1928.
[33] G. Pieri, L. Pusillo, Interval values for multicriteria cooperative games, AUCO Czech Economic Review 4 (2010) 144-155.
[34] G. Pieri, L. Pusillo, Multicriteria partial cooperative games, Applied Mathematics 6 (2015) 2125-2131.
[35] V. Podinovskii, V. Nogin, Pareto-optimal solutions of multicriteria problems, Nauka, 1982.(in Russian).
[36] J. Puerto, F. Perea, On minimax and Pareto optimal security payoffs in multicriteria games. Journal of Mathematical Analysis and Applications. 457(2) (2018) 1634—1648 https://doi.org/10.1016/j.jmaa.2017.01.002
[37] P. Reddy, E. Shevkoplyas, G. Zaccour, Time-consistent Shapley value for games played over event trees, Automatica 49 (6) (2013) 1521-1527.
[38] A. Rettieva, Cooperation in dynamic multicriteria games with random horizons, J Glob Optim (2018). https://doi.org/10.1007/s10898-018-0658-6.
[39] A. Sedakov, On the Strong Time Consistency of the Core, Automation and Remote Control, 79(4), (2018) 757-767.
[40] L. Shapley, A value for n-person games, Contributions to the Theory of Games, II, H. Kuhn and A. W. Tucker, eds., Princeton: Princeton University Press, (1953) 307-317.
[41] L. Shapley, Equilibrium points in games with vector payoffs, Naval Research Logistics Quarterly 6 (1959) 57-61.
[42] M. Voorneveld, D. Vermeulen, P. Borm, Axiomatizations of Pareto equilibria in multicriteria games, Games Econom. Behav. 28 (1999) 146–154.
[43] M. Voorneveld, S. Grahn, M. Dufwenberg, Ideal equilibria in noncooperative multicriteria games, Math. Methods Oper. Res. 52 (2000) 65–77.
[44] Y. Wang, Existence of a Pareto equilibrium, Journal of Optimization Theory and Applications, 79 (1993) 373–384.
[45] D. Yeung, An irrational-behavior-proof condition in cooperative differential games, International Game Theory Review 8 (4) (2006) 739–744.